



Just take a deep breath...

A review to compare the effects of spontaneous versus directed Valsalva pushing in the second stage of labour on maternal and fetal wellbeing

Andrea Bosomworth, Josette Bettany-Saltikov

Background: This work was undertaken as part of a Bachelors degree in Midwifery. The main focus of this article is around the methods used to review and synthesise the literature in order for it to be useful in practice. Management of the second stage of labour for women without epidural analgesia varies considerably. The Valsalva manoeuvre was introduced into obstetrics as a way of reducing the duration of the second stage in order to limit the exposure of the fetus to the perceived risks of this stage of labour.

Aims: This review aimed to compare the effects of directed Valsalva pushing with spontaneous pushing on: duration of the second stage, mode of delivery, perineal outcome, Apgar scores, cord blood pH and fetal heart rate.

Electronic resources and hand searching were undertaken for quantitative research carried out on women without epidural analgesia. No restrictions were made based on parity or gestation, however all included studies involved women at 37–42 weeks gestation.

Main results: Ten studies were included in the review – three randomised controlled trials (including one pilot study), five non-randomised controlled trials (including one pilot study), one retrospective analysis and one controlled trial using

intra-subject replication. Very few statistically significant results were reported and it would appear that directed Valsalva pushing has no effect on the mode of delivery. Although directed Valsalva pushing may shorten the second stage of labour, the results were inconclusive and contradictory. The outcome most affected by directed Valsalva pushing was perineal integrity, with significantly poorer outcomes being reported in terms of severity and frequency of perineal tears. In considering fetal wellbeing, no differences were found between directed Valsalva pushing and spontaneous pushing for Apgar scores or umbilical cord blood pH values. The effect of directed Valsalva pushing on the fetal heart rate is inconclusive but there is some evidence to suggest that it may increase risk of late decelerations.

Conclusions: The routine use of directed Valsalva pushing does not appear to confer any significant benefits or risks in terms of duration of the second stage, mode of delivery, Apgar scores or cord blood pH values. There is a significant trend towards poorer perineal outcomes when directed Valsalva pushing is used. Directed Valsalva pushing may be a cause of late decelerations in the fetal heart rate. Given its lack of benefits and potential risks, the practice of directing women to use Valsalva pushing should be discouraged.



© Bubbles/Angela Hampton

Introduction

It is often assumed that historical practices, as well as those carried out in traditional societies, are inferior to those carried out in the developed world. However, evidence is emerging which now challenges these assumptions and offers more support to some of these more traditional practices, particularly in relation to pregnancy and childbirth (Roberts & Woolley 1996). One example of this can be found in the behaviour of women in the second stage of labour. If left to respond instinctively to the urges they feel in the second stage of labour, most women would spontaneously bear-down (Beynon 1957). However, it

has become common in Western culture for women to be encouraged or even directed to push for as long and as hard as they can in the second stage (Petersen & Besuner 1997, O'Connell *et al* 2001). This review focuses on how second stage pushing efforts should be managed for women without epidural analgesia in an attempt to provide evidence to support or refute the use of directed pushing, or what is termed the Valsalva manoeuvre.

Definitions of terminology

Within this review, spontaneous pushing is defined as a woman responding purely to the urges of her body, ie in response to Ferguson's bearing down reflex. This reflex occurs as the presenting part descends below the ischial spines, stretching the nerves and muscles, stimulating the stretch receptors in the posterior vaginal wall causing an increase in the release of oxytocin (Petersen & Besuner 1997, Roberts 2002).

The action of taking a deep breath and holding it for a prolonged period whilst bearing down against a closed glottis is known as the Valsalva manoeuvre (Rushmer 1947). Directed Valsalva pushing within this review is defined as occurring when the midwife controls the woman's pushing by instructing her when and how to push. Typically, this means instructing the woman to take a deep breath at the onset of a contraction, holding it and pushing for 8 to 10 seconds against a closed glottis, aiming for three pushes per contraction.

Background

It is thought that one of the reasons for the introduction of directed Valsalva pushing was its ability to shorten the second stage of labour and therefore reduce the exposure of the fetus to the perceived risks of a prolonged labour (Rossi & Lindell 1986). These perceived risks were based largely on the work of Hellman and Prystowsky in 1952 (cited in Roberts & Woolley 1996, p417) who found a strong correlation between perinatal morbidity and mortality and a prolonged second stage. More recent studies have also linked prolonged second stage to fetal compromise (Neuberg 1995, Nordstrom *et al* 2001). As a result, midwives and doctors became concerned with finding ways of reducing the duration of the second stage, thereby reducing the frequency and severity of neonatal morbidity. Early trials on the use of directed Valsalva pushing indicated that it led to a shorter second stage (Barnett &

Humenick 1982) and it therefore grew in popularity as the safest way to manage the second stage of labour (Petersen & Besuner 1997, Roberts 2002). Ongoing research into the correlation between length of the second stage and fetal well-being, using indicators such as cord pH levels, Apgar scores and admissions to neonatal intensive care units, have not identified this at statistically significant levels even where the second stage has lasted up to four hours (Menticoglou *et al* 1995, Myles & Santolaya 2003).

When considering the physiological effect on the fetus, it has been suggested that strenuous and sustained pushing against a closed glottis leads to an increase in intra-thoracic pressure, which in turn reduces the venous return to the heart. A drop in blood pressure then occurs, which causes a fall in cardiac output and a subsequent reduction in the perfusion of oxygenated blood to the uterus, placenta and ultimately the fetus (Rushmer 1947), resulting in reduced pH levels, hypoxia and compromising changes in the fetal heart rate (Caldeyro-Barcia 1979, Knauth & Haloburdo 1986). However, other research has found no significant difference in fetal well-being between different methods of pushing (Barnett & Humenick 1982, Thomson 1993).

Another effect of prolonged directed Valsalva pushing is the increased risk of maternal fatigue, which potentially increases the likelihood of instrumental delivery (Mayberry *et al* 1999/2000). However, much of this evidence derives from trials in which women have had epidural analgesia, and it is therefore necessary to view the conclusions with an understanding that the same may not be true for women using other forms of analgesia. One of the aims of this review is to consider evidence from research based on women who are labouring without epidural analgesia.

There are times when directed pushing may be advantageous, for example if the woman is fearful of the second stage or has become particularly anxious (Chalk 2004a). It is also widely accepted that epidural analgesia necessitates a greater degree of direction from the midwife in view of the diminished sensation of the Ferguson's reflex (Mayberry *et al* 1999/2000, Hansen *et al* 2002). Although the involuntary need to bear down occurs spontaneously in the majority of labours (Roberts *et al* 1987), research also indicates that its occurrence cannot be assumed, even in women without epidural analgesia (McKay *et al* 1990). For

these women, having a carer who can direct their pushing efforts may be beneficial.

Since the 1970s, the weight of evidence and published literature has indicated that the use of directed Valsalva pushing should be discouraged on the basis that the potential harm it can cause outweighs its benefits (Parnell *et al* 1993). Despite this, it would appear that this technique is still advocated for use in the second stage (Aldrich *et al* 1995, Petersen & Besuner 1997, Chalk 2004b). It is possibly the contradictory nature of the evidence and the lack of systematic review that has led to conflicting practice and, in light of this, there is a need to re-examine the evidence for directed Valsalva pushing to consider whether it reduces the duration of the second stage enough to make a difference to fetal outcomes.

Objectives of the review

The review will evaluate existing research to consider the impact of spontaneous or directed Valsalva pushing on the outcomes of the second stage of labour in women without epidural anaesthesia. The outcomes compared will be: fetal/neonatal well-being, measured by heart rate changes, cord blood pH levels and Apgar scores at birth; the duration of the second stage of labour; mode of delivery; and the presence and degree of perineal trauma.

Search strategy

Fifteen international databases were searched as part of this review, these included the Cochrane database, MEDLINE, CINAHL, MIDIRS and ZETOC. Keywords used for searching included 'labour', ('labor'), 'second stage management', 'pushing', 'bearing-down', 'Valsalva', 'spontaneous' and 'directed'.

In addition to keyword searching, author searches were carried out on the above databases using authors identified within the first selection process. Reference lists were also searched for further relevant trials and studies, including those on NHS Trust guidelines and protocols relevant to the second stage of labour and those in reputable midwifery texts. Further searching was carried out using generic internet search engines and by hand searching of texts and journals.



Inclusion criteria

Types of studies

Full text reports of randomised controlled trials (RCTs), controlled trials (CTs) and retrospective quantitative studies were included in the selection process, but trials involving women with epidural analgesia were excluded. Qualitative studies were not included in the selection process. As very little research has been carried out recently on women who did not have epidural analgesia, there was no restriction on the age of the studies which were to be considered.

Participants

The review included studies for both primigravida and multigravida women in the second stage of labour with singleton, cephalic pregnancies, uncomplicated by any obstetric or other medical condition. No other specific inclusion criteria applied.

Interventions

Included studies examine the effects of directed Valsalva pushing and/or spontaneous pushing on outcomes related to maternal and fetal well-being.

Outcomes

These were identified as measurement and a record of at least one of the following outcomes: duration of the second stage, mode of delivery, cardiotocographic data, cord blood pH levels, Apgar score and perineal outcome.

Method of review

Selection of studies

Academic requirements necessitated the inclusion of ten primary research papers in the review. Two levels of selection were carried out by two independent reviewers using a structured and tested checklist. Only those studies identified for inclusion by both reviewers were included in the review.

Review of methodological quality

The purpose of reviewing the quality of the studies was to ascertain the degree of methodological rigour which would enable the results to have internal and external validity and reliability. As all the included studies were based on quantitative methodologies, the McMaster University Critical

Review framework was used to undertake this (Law *et al* 1998). A single reviewer carried out the review for quality. It is acknowledged that this may have introduced bias into the process; however, it was considered that this was minimised by the use of a well-constructed review form.

Data extraction

Data were extracted on the effects of spontaneous and directed Valsalva pushing in the second stage of labour on the woman and the neonate. Details were also extracted about the populations and any specific interventions used in order to consider the comparability of the studies in terms of demographic and pregnancy details. Finally, the results presented in each study relating specifically to the outcomes of this review were recorded on a standard form to minimise the risk of bias.

Description of the studies

A total of 673 titles were identified from the initial search; 590 were rejected as they were found not to fit the criteria, and a further 60 were of relevance to the topic but were non-research articles. This left 23 research papers for the stage 1 selection process. Sixteen papers were then selected for second stage selection, from which 10 quantitative studies were identified for inclusion in the review (the number required by the university). Table 1 presents a summary of the included studies.

Summary of the review of methodological quality

Use of the McMaster Critical Review framework (Law *et al* 1998) identified that all the studies selected for the review had factors which would limit the validity of their results. Using the framework assessment for quality, the overall rating of

the methodological quality of each paper identified two papers as 'good' – Paine and Tinker (1992), Yeates and Roberts (1984). Six papers were classed as 'moderate'. Two papers were classed as 'poor' – Perry and Porter (1979), Parnell *et al* (1993). As previously stated, academic requirements necessitated inclusion of ten primary research papers in the review and it was therefore necessary to include the results from these two papers although they had been identified as being methodologically poor.

Results

In comparing outcomes for the ten studies, the variable quality of the data needs to be acknowledged but the overall process of review was undertaken as a systematic process allowing the maximum information to be extracted for each study.

Duration of the second stage of labour

This outcome was reported by nine studies (graph 1). In the study by Parnell *et al* (1993), the groups were reassigned part way through the trial as it was found that the descriptions 'open glottis' and 'closed glottis' more accurately defined how the women were actually pushing. Parnell *et al* (1993) redefined the groups as 'closed' or 'open' glottis when it became apparent that this better defined the way women were actually pushing. Results are the presented for these new groups.

Mode of delivery

There was no significant difference between the two groups in terms of mode of delivery, according to the three papers that reported this outcome. Thomson (1993) found that 20% of the

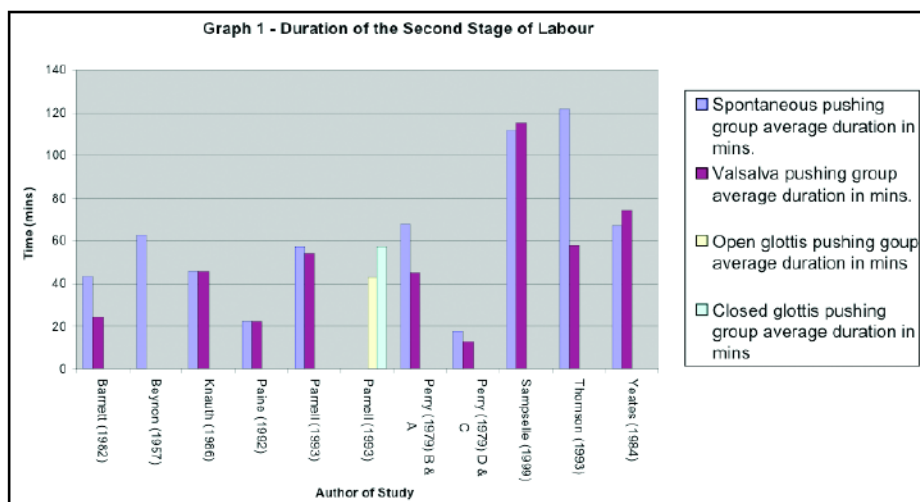


Table 1. Summary of included research papers

Study	Population	Intervention	Control	Outcomes	Results relevant to review
Barnett & Humenick (1932)	Low risk, multigravida 16–30 years 38–42 weeks gestation	Spontaneous pushing with urges	Valsalva pushing technique	Fetal heart rate Cord blood pH Duration of 2nd stage	Intervention group longer labour, higher cord arterial pH both non-significant. No significant difference in Apgar scores
Beynon (1957)	'Normal' Primigravida Vertex presentation	Spontaneous pushing with urges	Valsalva pushing technique	Duration of 2nd stage Birth weight Mode of delivery Perineal outcome	Intervention group 50% fewer forceps delivery, no comparison of duration. Lower suture rate for intervention group no statistical analysis
Knauth & Haliburdo (1936)	No details	Exhalatory pushing technique in response to urges	Breath holding pushing	Duration of 2nd stage Maternal stress Fetal stress Pain perception	No differences in duration. 30% control group vs 18% intervention group severe fetal variable decelerations. 30% control group vs 54% intervention group maintained average fetal heart baseline variability
Paine & Tinker (1992)	Uncomplicated pregnancy. 38–42 weeks gestation Singleton, OA position No fetal risks, FH 120–160 bpm	Spontaneous pushing with urges	Valsalva pushing technique	Cord blood pH Duration of 2nd stage	No significant difference between groups in terms of cord pH or duration of 2nd stage
Parnell <i>et al</i> (1993)	Primigravida Singleton Vertex presentation 37+ weeks gestation	Spontaneous pushing with urges	Valsalva pushing technique	Cord blood pH Duration of 2nd stage Apgar score Perineal outcome Mode of delivery Augmentation rate	No significant difference between original groups in terms of cord pH or duration of 2nd stage. Comparisons for groups based on open or closed glottis indicated shorter 2nd stage for open glottis group. No significant difference in mode of delivery
Perry & Porter (1979)	Primigravida and multigravida	Group B & D Untrained in pushing, spontaneous to urges	Group A & C Diaphragmatic, Valsalva style pushing taught in classes	Duration of 2nd stage	Women using taught breath-holding pushing had shorter 2nd stage, particularly primigravida 45mins vs 68 mins. Multigravida times 13 mins control vs 18 mins intervention
Sampsel & Hines (1999)	Primigravida 18+ yrs No complications	Spontaneous pushing with urges	Directed in Valsalva pushing technique	Duration of 2nd stage Perineal outcome Pain perception in puerperium	Spontaneous pushing resulted in more intact perineum, fewer episiotomies and tears. Spontaneous pushing mean duration 2nd stage 111.6 mins vs 114.9 mins for directed pushing
Thomson (1993)	Primigravida 18+ yrs Singleton, cephalic 37+ weeks gestation No complications	Spontaneous pushing with urges	Valsalva pushing technique	Cord blood pH Resuscitation needs Duration of 2nd stage Perineal outcome Mode of delivery	No differences in mode of delivery, perineal outcome, cord blood pH. Intervention group 2nd stage duration significantly longer than control group, 121.4 mins vs 58 mins.
Woolley & Roberts (1995)	Primigravida 37–42 weeks gestation No complications	Mini pushing with open glottis	Valsalva pushing technique	Intrauterine pressure Fetal heart rate	Some response of the FH to different pushing not significant other than that Valsalva pushing resulted in deeper decelerations than mini-pushing
Yeates & Roberts (1984)	Primigravida 38–42 weeks gestation No complications Spontaneous onset of labour, OA position	Instructed how to co-ordinate pushing efforts with body's urges	Valsalva pushing technique	Duration of 2nd stage Apgar score Perineal outcome Maternal effort	No significant difference in duration, Apgar score. Significant difference in perineal outcome, fewer and less severe in intervention group.

spontaneous pushing group and 12 % of the Valsalva pushing group required instrumental delivery ($\chi^2 = 0.0232$, $P = 0.8785$).

Perineal outcome

The results reported for this outcome demonstrated a trend towards less frequent and less severe perineal trauma for women in the spontaneous pushing groups. Beynon (1957) reported that 39% of women using spontaneous pushing required suturing compared with 63% of women in the Valsalva pushing group. Thomson (1993) also measured perineal outcome in terms of the need for suturing, stating that 73.3% of women using

spontaneous pushing, compared to 58.8% of women using Valsalva pushing, required suturing of the perineum ($\chi^2 = 0.2396$, $P = 0.6305$). Paine and Tinker (1992) reported no significant differences between the two groups in terms of the number of women requiring suturing.

In three other studies, perineal outcome is classified into categories based on the severity of tear, whether episiotomy was required or whether the perineum was intact after delivery. Parnell *et al* (1993) found no statistically significant differences between the closed and open glottis groups although a higher number of women required episiotomies in the closed glottis group ($\chi^2 = 3.95$, $P < 0.05$). Sampsel and Hines (1999) reported

significant differences ($\chi^2 = 8.1$, $P = 0.043$) between the two groups: 45.45% of women had intact perineums compared to only 7.14% in the Valsalva pushing group. Yeates and Roberts (1984) documented similar results, with 80% of the spontaneous pushing group having intact or first degree tears compared only 20% of the Valsalva pushing group achieving this outcome ($P < 0.05$).

Fetal and neonatal well-being

Six out of the ten studies recorded measurements for this outcome although the variable chosen as the measurement tool was not the same in each of the six papers. The measurements to be used as



comparisons in this review are fetal heart rate, specifically the frequency of late or variable decelerations in and the maintenance of an average baseline that is within accepted satisfactory range, ie 120–160 bpm (Bennett & Brown 1999). In terms of the neonate, the measurement variables compared are the Apgar scores recorded for one and five minutes and the umbilical cord arterial blood pH value.

Fetal heart rate

Three studies considered the effect of pushing on fetal heart rate. Woolley and Roberts (1995) found no significant difference in the rate of late or severe variable decelerations between the two groups ($\chi^2 = 1.9286$, $P = 0.1649$). Similarly, there was no significant difference in the number of those who maintained a normal baseline fetal heart rate. In the study by Barnett and Humenick (1982), there were no late or severe variable fetal heart rate decelerations in the spontaneous pushing group compared with two such observations in the Valsalva pushing group. These results were not tested for statistical significance. The results from the study by Knauth and Haloburdo (1986) showed that 30% of fetuses in the Valsalva pushing group experienced severe variable decelerations and 30% maintained average baseline variability. This compares to 17.6% of fetuses in the spontaneous pushing group experiencing severe variable decelerations and 58.8% maintaining average baseline variability (no statistical analysis was undertaken on these data).

Apgar scores

One and five minute Apgar scores were recorded by Barnett and Humenick (1982), Paine and Tinker (1992) and Parnell *et al* (1993). Each found the average Apgar score at one minute and five minutes to be the same in the spontaneous pushing and Valsalva groups.

Umbilical cord blood pH

Four studies measured the umbilical cord blood pH following delivery and found no statistically significant differences in the values between the spontaneous pushing group and the Valsalva pushing group. Paine and Tinker (1992) ($P = 0.53$), Barnett and Humenick (1982) ($P = 0.05$), Thomson (1993) ($P = 0.184$) and Parnell *et al* (1993). Thomson (1993) found a negative correlation between the length of the second stage and the umbilical cord blood pH value in the Valsalva pushing group but the same correlation did not

exist in the spontaneous pushing group. Parnell *et al* (1993) found no statistically significant differences between the two groups in terms of Apgar scores or umbilical cord blood pH values; however, all eight neonates with cord blood pH values <7 and/or Apgar scores of <7 were born to women using a closed glottis pushing technique.

Discussion

From the ten papers included in this review, very few statistically significant differences were identified between spontaneous and directed Valsalva methods of pushing for the outcomes being considered.

Duration of second stage

Parnell *et al* (1993) found a significantly shorter second stage for women with an open glottis during pushing when compared to women with a closed glottis. Sampsel and Hines (1999) and Yeates and Roberts (1984) reported a similar though non-significant trend with spontaneous pushing groups having slightly shorter second stages of labour than the Valsalva groups. However, Thomson (1993) found contradictory results, reporting a significantly shorter second stage for the Valsalva pushing group. These results are similar to those of Barnett and Humenick (1982) and Perry and Porter (1979) who do not report statistical significance.

Considering mode of delivery, Beynon (1957) reported a 50% reduction in the rate of forceps deliveries in the spontaneous pushing group. However, Parnell *et al* (1993) and Thomson (1993) found no significant difference in the rate of instrumental deliveries between the control and intervention groups.

The outcome that generated the most statistically significant results was the frequency and severity of perineal trauma. Beynon (1957) reported a 24% lower suture rate in the spontaneous pushing group than in the directed Valsalva pushing group. Sampsel and Hines (1999) and Yeates and Roberts (1984) also found less perineal trauma in the spontaneous pushing groups, reporting significantly fewer and less severe lacerations in their spontaneous pushing groups. Thomson (1993), however, found a slightly higher number of women requiring suturing in the spontaneous pushing group although the difference

between the groups was not statistically significant.

None of the six papers that considered fetal well-being as an outcome reported any statistically significant differences between the control and intervention groups in terms of fetal heart rate. Barnett and Humenick (1982) and Knauth and Haloburdo (1986) both reported a non-significant trend towards fewer late and severe decelerations associated with spontaneous pushing. The five studies considering umbilical cord arterial blood pH and Apgar scores found no significant differences between the two methods of pushing.

Included studies

There was extensive searching of the literature but only one of the papers included in the review was published in the last five years. For the remainder, the majority were published within the last 15 years and the oldest study dated back to 1957 (Beynon 1957). In many situations it would be preferable, even essential, that research used as a basis for practice is more recent than this, yet, as Cutcliffe and Ward (2003) point out, it is difficult to identify a specific age when a piece of research becomes too old to be relevant and in this context, when considering midwifery practice and the measurement of standard outcomes and the paucity of literature, to include all of the studies over this time frame was a pragmatic decision that appeared unlikely to prejudice the review.

The ten studies included in the review represent research into the topic carried out over a wide time span with an international perspective in that they include research from the USA, Canada, UK and Denmark. This offers a degree of external validity; however, due to the differences in maternity care practices around the world, the use of international research could reduce the comparability of findings. Although randomised controlled trials are often cited as the gold standard for quantitative research, in midwifery there is a need to address the holistic aspects and implications of practice, some aspects of which are not conducive to measurement within the structure of a randomised controlled trial (Clark 2000). The inclusion of five trials within this review that were not randomised could create a potential weakness in the results. Similarly, the inclusion of a retrospective study (Sampsel & Hines 1999), where the participants were between nine and

14 months post-delivery, is also contentious due to possible recall bias as some women may have forgotten specific details about their labour (Parahoo 1997). Also potentially significant in terms of the quality of the results is the inclusion of two pilot studies (Thomson 1993, Yeates & Roberts 1984).

Review of quality of included studies

Sampling and groups

None of the studies described why they had chosen a particular sample size. Only two studies presented results for more than 40 women (Beynon 1957, Parnell *et al* 1993), and two studies included only ten women (Barnett & Humenick 1982, Yeates & Roberts 1984). A small sample size reduces the external validity of the results of a trial (Polit & Hungler 1995); therefore, taking this into account, the results of studies by Barnett and Humenick (1982) and Yeates and Roberts (1984) must be viewed with considerable caution. In addition to sample size, the method of sampling and the resulting degree of representativeness of the sample to the population is important in terms of external validity of the results (Parahoo 1997). Perry and Porter (1979) and Thomson (1993) do not indicate how they carried out the sampling process. The other eight studies used either convenience or volunteer sampling or a combination of both methods. Both convenience and volunteer sampling introduce weakness into a trial – they are unlikely to yield representative samples and the results from these trials will have a poor level of external validity (Argyrous 2000).

It is acknowledged that the inclusion of papers using a mixture of both primigravida and multigravida women, or indeed papers which do not specify the parity of the women, limits the internal validity and comparability of the studies. However, following further scrutiny of the data, excluding these papers does not affect the overall findings of this review.

Of the studies included in the review, only two explicitly state that subjects were randomised to the control or intervention groups (Knauth & Haloburdo 1986, Parnell *et al* 1993). There is therefore a risk of systematic error in the other eight trials. Only four studies tested the groups for comparability before commencing (Yeates & Roberts 1984, Paine & Tinker 1992, Parnell *et al* 1993, Thomson 1993). In other words, the groups in the trials by

Barnett and Humenick (1982), Beynon (1957) and Perry and Porter (1979) were not randomly allocated and were not tested for comparability. It is therefore not possible to assume that the results measured were attributed solely to the intervention being studied. Thomson (1993) tested for matching of groups and found them to be comparable on all factors other than duration of the first stage and use of pethidine. The spontaneous pushing group was found to have a significantly shorter average length of the first stage of labour and a marginally non-significant greater use of pethidine (Thompson 1993). The positive correlation between both of these factors and the duration of the second stage of labour (Thomson 1993) makes it difficult to ascertain whether the differences between the two groups occurred because of the intervention or because of these confounding factors.

Although randomisation is preferable in experimental design, within healthcare it is not always possible or ethical. The subject of this review provides a good example of the inappropriateness of RCTs for some research. It could be argued that it is not ethically acceptable to insist that a woman pushes in a particular way during labour and that it is more appropriate for her to be given the choice of whether she wishes to use a directed Valsalva or spontaneous method of pushing, as in the study by Paine and Tinker (1992). However, this also requires more flexibility in our understanding of the available evidence upon which to base our practice.

Intervention and control

The controlled trials and the single subject experiment describe the intervention and the control in some detail. The intervention for each study was a variation on encouraging women to push in response to their bodies' natural urges. Knauth and Haloburdo (1986), Paine and Tinker (1992) and Woolley and Roberts (1995) were slightly more prescriptive in how they asked women in the intervention group to push and it could be argued that this does not truly constitute spontaneous pushing. Perry and Porter (1979) make the assumption that because women did not attend formal antenatal teaching on pushing techniques that they would push in a spontaneous manner. This fails to take account of any other coaching the women may have received, thus reducing the internal validity of the results. In contrast, the control group pushing method for each of the studies was described and involved the women being directed by the

carer to use the Valsalva manoeuvre during pushing. This included asking the women to take a deep breath and hold it whilst bearing down against a closed glottis. Four papers specifically asked women to sustain effort for at least 10 seconds whilst the other requested that women push for as long as possible. Taking these points into account, the interventions and controls used in these nine trials are reasonably comparable. The question relating to pushing used by Sampsel and Hines (1989) in their retrospective study was developed by experienced practitioners in order to increase internal validity. However, the description used for directed pushing is vague, "told how and when to push by someone" (Sampsel & Hines 1999, p37). This does not necessarily imply that they were told to push using a particular technique and therefore their conclusions about the effects of specific pushing methods in this study are more questionable.

Three of the studies reported contamination between the intervention and the control groups due to subject error. The results of the trial by Thomson (1993) need to be viewed with caution since nearly a quarter of the women in the control group did not use the correct pushing technique. Parnell *et al* (1993) also found that women were unable to use the techniques originally described to them for either spontaneous or directed Valsalva pushing. This issue was addressed by reassigning the groups during the trial depending on the technique that they were actually using, to either an open glottis group (intervention group) or closed glottis group (control group). Although doing this reduced the degree of subject error (Clarke & Croft 1998), it also reduces the degree of comparability of the results of this trial to the others within the review. By using a single subject with intra-subject replication, Woolley and Roberts (1995) avoided the risks of confounding factors since each woman acted as her own control. However, as emphasised by Gomm *et al* (2000), the risk of this design is that the effects of the intervention may still be present when the woman switches to the control. The other seven studies do not report subject error. This may be because the subjects were able to successfully adhere to the prescribed pushing method. However, given the experiences of Thomson (1993) and Parnell *et al* (1993), it may be instead that the other studies failed to take account of this potential experimental weakness in reporting their results.

Barnett and Humenick (1982), Knauth and Haloburdo (1986), Thomson (1993) and Yeates and Roberts (1984) avoided observer/recorder error by using only one midwife/nurse to provide care to all of the women in their trials, thus increasing the reliability of their results. In contrast to this, all midwives working within the hospital were involved in the trial by Parnell *et al* (1993), despite a documented lack of interest in the trial by some of them. A lack of commitment by those being asked to take part in a trial can introduce observer error either consciously through the falsification of results, or subconsciously through the effect of observer bias (Parahoo 1997). Since independent observers were not used in the trial, this weakens the reliability of the results.

Outcomes and measures

Duration

Eight of the studies reported this outcome, but only Knauth and Haloburdo (1986) provided a definition. Using the duration of the second stage as an outcome presents a significant problem in that it makes the assumption that it is possible to identify both a start and an end time. The end of the second stage is generally accepted to be the time of complete delivery of the baby and an accurate time can be recorded for this. The start time is traditionally equated to the time when the cervix becomes fully dilated (Roberts 2002). However, without frequent internal examinations, it is not possible to accurately ascertain when this is. Therefore, a considerable degree of inaccuracy is associated with the calculation of the duration of the second stage, making it a poor indicator of the effects of different pushing techniques. A better outcome may be the actual time spent pushing in the second stage. This could be measured from the time that the first push occurred, whether that is spontaneously in response to Ferguson's reflex or as a result of direction from the carer.

Only Thomson (1993) and Parnell *et al* (1993) report a significant difference in duration between the groups. As Thomson's results may have been affected by confounding factors, only one study identified that open glottis pushing may have a shortening effect on the second stage of labour (Parnell *et al* 1993). Further high quality research, using a more accurate measurement tool, is necessary to confirm or reject this finding.

Mode of delivery

For the purpose of this review, the categories used are 'unassisted delivery' or 'instrumental delivery'. Mode of delivery is considered to be methodologically acceptable as an outcome; however, it is not completely robust. It was not possible to blind the carers to the intervention being applied and it can therefore be argued that their own bias could have influenced their decisions about how to proceed with delivery, particularly in the case of prolonged second stage. The introduction of this type of bias could have a significant impact on the accuracy of the results (Schultz *et al* 1995). In the studies by Beynon (1957) and Thomson (1993), it is also possible that inter-carer errors may have occurred since more than one midwife was involved in the study and no indication is given about whether they were instructed in how to decide when an instrumental delivery was indicated. These limitations may explain why the results of the three studies investigating the impact of pushing method on mode of delivery found contradictory results. The effect of changing attitudes towards medical intervention may also have a significant effect on the comparability of the results. In the 1950s, intervention in childbirth was almost routine (Drife 2002). However, by the 1990s, the emphasis in maternity care had refocused on the facilitation of a normal childbirth experience for women, with minimum intervention (DoH 1993). Given the previously discussed limitation of both of these trials, it is not possible to gauge the impact of pushing technique on mode of delivery based on the findings of this review.

Perineal outcome

Perineal outcome provided the most statistically significant results. On initial consideration the use of perineal trauma as an outcome would appear to be experimentally appropriate as it has a clear classification for measurement, ie intact, first, second and third degree tears or episiotomy. The reliability of the results within each trial and the comparability of results across the studies does however rely on inter-observer reliability. None of the three studies considering this outcome provides definitions of perineal trauma (Yeates & Roberts 1984, Thomson 1993, Sampsel & Hines 1999), so it is likely that inter-observer errors will have affected the results. The use of frequency of suturing as a measure of perineal outcome (Beynon 1957) is also questionable unless clear guidelines are included as to what constitutes a need for suturing. Gomm

et al (2000) state that the amount of inter-observer error is considerable when measuring subjective outcomes in healthcare. It is therefore necessary to view Beynon's (1957) results in relation to perineal outcome with caution. Overall, this review would indicate that spontaneous pushing is preferable to directed Valsalva pushing in terms of the frequency and severity of perineal trauma.

Fetal/neonatal well-being

This outcome was chosen for inclusion in the review as there appeared to be a widespread view that one of the most deleterious effects of directed Valsalva pushing is its effect on the fetus. Research by Caldeyro-Barcia (1979) and Bassell *et al* (1980) states that prolonged, strenuous bearing down against a closed glottis potentially causes late decelerations in the fetal heart rate, fetal hypoxia and fetal acidosis. Although often cited within articles, the aforementioned research is methodologically weak and the external validity of the results is low given the very small sample sizes involved.

Despite being a valuable outcome, it is not easy to measure, particularly when considering the fetus rather than the neonate. The presence of late decelerations in the fetal heart rate and changes in the baseline are routinely used as indicators of fetal compromise during labour (NICE 2001), and are therefore considered to be suitable for use as measures within this review. Reliability of the results was high in the studies by Woolley and Roberts (1995) and Barnett and Humenick (1982), as internal fetal scalp electrodes were used to collect data. Knauth and Haloburdo (1986) used an abdominal transducer to record the fetal heart rate. This may have introduced an element of recorder error into the results since there is some evidence that their use in the second stage does not accurately record fetal heart rate, particularly during pushing efforts (Kuhnert & Schmidt 2004). The results of two of the studies concur with the findings of previous research in suggesting that directed Valsalva pushing appears to have a deleterious effect on the fetus as indicated by late decelerations in the fetal heart rate and an inability to maintain a reassuring baseline rate (Barnett & Humenick 1982, Knauth & Haloburdo 1986). However, neither of these results was tested for statistical significance and, given the small sample sizes involved in these two studies, it is difficult to draw scientific conclusions from their results. Woolley and Roberts (1995) did test for statistical

significance and found no difference between the intervention and control pushing methods. Another potentially important finding in the study by Woolley and Roberts (1995) was that the fetuses who experienced decelerations in their heart rates with one type of pushing were very likely to experience them with both methods. The significance of this finding is that it may imply that fetuses who are compromised by pushing during the second stage would be compromised irrespective of the method of pushing used by the woman. The lack of any other research using intra-subject replication prevents this hypothesis being either substantiated or disputed. In considering the importance of these results, it is necessary to take account of the potential weaknesses created by the use of intra-subject replication as an experiment design. Woolley and Roberts (1995) themselves highlight the potential problem stating that 'washout' may well have caused the effects of one method of pushing to influence the results recorded for the alternative method. The results of this review suggest that directed Valsalva pushing may have a deleterious effect on the fetus as measured by changes in the fetal heart rate, but larger, high quality trials using advanced methods of fetal heart rate monitoring and statistical analysis of results are necessary to substantiate this further.

Considering the use of umbilical cord arterial blood pH level as a measure of neonatal well-being, assuming that the samples are drawn immediately following delivery, should mean that such measurement is likely to be reliable (Westgate *et al* 1994). All four studies state this to have been the case. However, comparability between studies may be difficult due to the use of different equipment within the four studies. This may explain why the values recorded by Parnell *et al* (1993) are all noticeably lower than those recorded by Thomson (1993) and Barnett and Humenick (1982). The fact that cord blood pH values were not significantly different for the two methods of pushing is interesting as it disputes the commonly held view, based on Caldeyro-Barcia's research (1979, 1981), that directed Valsalva pushing causes fetal acidosis. Thomson (1993) and Parnell *et al* (1993) actually found a slight trend towards higher pH levels for the directed Valsalva pushing groups. These findings agree with those of Wood *et al* (1973); however, in this study, the pH levels appear to correlate strongly with the length of the second stage and this is not taken into account in the analysis of the

results. One of the most significant issues with these results is the effect of confounding factors on fetal acidosis. Considering first duration, Thomson (1993) found a statistically significant negative correlation between the duration of the second stage and the pH levels of neonates in the directed Valsalva pushing group. However, the size of this study means that the external validity of the result is low. The study by Barnett and Humenick (1982) attempts to analyse the effects of other factors on cord blood pH levels, namely frequency, strength and duration of contractions, and reports a negative correlation between the frequency of contractions and cord blood pH levels at delivery.

Overall, since none of the studies reported any statistically significant difference between the pushing methods in terms of cord arterial blood pH level at delivery, it is not possible to determine from these results whether either method is more deleterious to the neonate than the other. The fact that directed Valsalva pushing has a significant negative correlation with cord blood pH levels in prolonged second stage of labour is important (Thomson 1993), although difficult to apply in practice since it is impossible to predict the duration of the second stage.

A more accurate measure of fetal well-being might be cord blood base excess values. The reason for this is that, although cord blood pH does indicate acidosis within the neonate, it does not differentiate between respiratory acidosis and metabolic acidosis (Westgate *et al* 1994, Mason *et al* 2001). Respiratory acidosis is caused by a build up of carbon dioxide in the fetal blood, which is often a harmless result of cord compression in the second stage of labour and is not necessarily indicative of fetal compromise (Mason & Paterson-Brown 2001). Metabolic acidosis occurs when there is a build up of lactic acid due to the fetus having to use anaerobic rather than aerobic metabolism; this is usually a result of serious hypoxia during labour (Rooth 1988, Mason & Paterson-Brown 2001) and its identification is indicative of fetal compromise. Base excess values provide a way of differentiating between acidosis that is respiratory and metabolic, and therefore they provide a better measure of fetal compromise than pH levels.

The final measure used to assess fetal well-being was the Apgar score. Although Apgar scores are calculated following a prescribed ordinal data table (Apgar 1953), it is recognised that there is considerable potential for subjectivity in

the scoring, creating poor inter-recorder reliability (Bharti & Bharti 2005). In order to be reliable as a measurement tool, blinding of the recorder to the treatment being used would be essential to avoid bias. None of the studies reporting Apgar scores were able to use blinding within their methodology and the results must therefore be viewed with caution. The internal validity of the Apgar score as a measure of neonatal well-being is also questionable. Research by Menticoglou *et al* (1995) suggests that the specificity of the Apgar score is low. The studies within this review failed to take account of this issue. The possibility of correlation between Apgar scores and other variables, such as duration of the second stage (Menticoglou *et al* 1995) or mode of delivery, was also not considered. The three studies failed to identify any difference between directed Valsalva pushing and spontaneous pushing in resultant Apgar scores (Barnett & Humenick 1982, Paine & Tinker 1992, Parnell *et al* 1993). However, given the lack of blinding and analysis of other correlations, these results have little clinical value and further evidence is required from appropriately designed trials in order to assess the effect of methods of pushing on the Apgar score.

Conclusion

Assuming that epidural analgesia is not being used, the urge to push as stimulated by Ferguson's reflex occurs naturally for the majority of women once fetal position and station are optimal. If they are allowed to respond to this urge, women will bear down spontaneously during contractions. Directed Valsalva pushing was introduced as a means of managing the second stage for wholly altruistic reasons. It was perceived to have considerable benefits for the woman and fetus by providing a means of shortening the second stage of labour, thus reducing the risks presumed to be associated with a prolonged second stage. Since directed Valsalva pushing represents a shift in control from the woman to the carer and effectively requires the woman to ignore her body's natural response to labour, there is an ethical and professional obligation to ensure that scientific evidence supports its use.

This review has identified that there has been very little significant research carried out to evaluate the effects of directed Valsalva pushing in the second stage of labour, particularly in the last decade. Having rigorously reviewed the methodological quality of existing

research, the authors feel justified in saying that the quality of the research available is generally not of a high enough standard to afford the results any clinical significance.

Careful consideration needs to be given to the outcomes and measurement tools used to assess the effects of directed Valsalva pushing. In particular, a more accurate method of timing the duration of the second stage is needed before it can be relied upon as an outcome. Research indicates that the Apgar score can provide an accurate measure of neonatal condition at birth, but only if blinding of the recorder to the treatment is part of the methodological design. Whilst umbilical cord arterial blood pH values can identify fetal acidosis, they do not differentiate between respiratory and metabolic acidosis, and base excess may be a better measurement tool.

Overall, the results appear to indicate no significant benefits or risks in using the directed Valsalva manoeuvre in terms of duration of the second stage, mode of delivery, Apgar score and cord blood pH values. Directed Valsalva pushing may be responsible for causing late decelerations in the fetal heart rate, though these results were not tested for statistical significance. The most significant finding from this review is the deleterious effect that directed Valsalva pushing has on perineal outcome, causing more frequent and more severe tears to occur than if women are pushing spontaneously. Given the issue of poor methodological quality, all of these results are largely inconclusive.

The authors conclude that, based on the findings of this review, for women without epidural analgesia who are experiencing Ferguson's reflex to bear down, directed Valsalva pushing should not be encouraged. There is a need for further high quality research to be conducted using large samples, reliable measures and taking account of the complexities of confounding factors and biases. Until this research is available, midwives should support women in their instinctive response to the second stage of labour.

References

Aldrich C, D'Antona D, Spencer J *et al* (1995). The effect of maternal pushing on fetal cerebral oxygenation and blood volume during the second stage of labour. *British Journal of Obstetrics and Gynaecology* 102(6):448-53.

Apgar V (1953). A proposal for a new method of evaluation of the newborn infant. *Current Research in Anesthesia and Analgesia* 32:260-7.

Argyrous G (2000). *Statistics for social and health research*. London: Sage Publications.

Barnett M, Humenick S (1982). Infant outcome in relation to second stage labor pushing method. *Birth* 9(4):221-9.

Bassell G, Humayun S, Marx G (1980). Maternal bearing down efforts – another fetal risk? *Obstetrics and Gynecology* 56(1):39-41.

Bennett VR, Brown LK *eds* (1999). *Myles textbook for midwives*. 13th ed. Edinburgh: Churchill Livingstone.

Beynon C (1957). The normal second stage of labour: a plea for reform in its conduct. *Journal of Obstetrics and Gynaecology* 64(6):815-20.

Bharti B, Bharti S (2005). A review of the Apgar score indicated that contextualization was required within the contemporary perinatal and neonatal care framework in different settings. *Journal of Clinical Epidemiology* 58(2):121-9.

Caldeyro-Barcia R (1979). The influence of maternal bearing down efforts during the second stage on fetal wellbeing. *Birth and the Family Journal* 6(1):17-21.

Caldeyro-Barcia R, Giusti G, Storch E *et al* (1981). The bearing-down efforts and their effects on fetal heart rate, oxygenation and acid base balance. *Journal of Perinatal Medicine* 9(suppl 1):63-7.

Chalk A (2004a). Pushing in the second stage of labour: part 1. *British Journal of Midwifery* 12(8):502-8.

Chalk A (2004b). Spontaneous versus directed pushing. *British Journal of Midwifery* 12(10):626-30.

Clark E (2000). The historical context of research in midwifery. In: Proctor S, Renfrew M *eds*. *Linking research and practice in midwifery*. Edinburgh: Bailliere Tindall: 35-54.

Clarke R, Croft P (1998). *Critical reading for the reflective practitioner: a guide for primary care*. Oxford: Butterworth-Heinemann.

Cutcliffe J, Ward M (2003). *Critiquing nursing research*. London: Quay Books.

Drife J (2002). The start of life: a history of obstetrics. *Postgraduate Medical Journal* 78(919):311-5.

Department of Health (1993). *Changing childbirth: report of the Expert Maternity Group*. London: DoH.

Gomm R, Needham G, Bullman A *eds* (2000). *Evaluating research in health and social care*. London: Sage Publications.

Hansen S, Clark S, Foster J (2002). Active pushing versus passive fetal descent in the second stage of labor: a randomized controlled trial. *Obstetrics and Gynecology* 99(1):29-34.

Knauth D, Haloburdo E (1986). Effect of pushing techniques in birthing chair on length of second stage of labour. *Nursing Research* 35(1):49-51.

Kuhnert M, Schmidt S (2004). Intrapartum management of non-reassuring fetal heart rate patterns: a randomized controlled trial of fetal pulse oximetry. *American Journal of Obstetrics and Gynecology* 191(6):1989-95.

Law M, Stewart D, Pollock N *et al* (1998). *Critical review form – quantitative studies*. McMaster University Occupational Therapy Evidence-Based Practice Group.

Mason R, Paterson-Brown S (2001). Fetal distress and blood lactate monitoring. *British Journal of Midwifery* 9(8):507-9.

Mayberry L, Wood S, Strange L *et al* (1999/2000). Managing second stage labor: exploring the variables during the second stage. *AWHONN Lifelines* 3(6):28-34.

McKay S, Barrows T, Roberts J (1990). Women's views of second stage labor as assessed by interviews and videotapes. *Birth* 17(4):192-8.

Menticoglou S, Manning F, Harman C *et al* (1995). Perinatal outcome in relation to second-stage duration. *American Journal of Obstetrics and Gynecology* 173(3 part 1):906-12.

Myles T, Santolaya J (2003). Maternal and neonatal outcomes in patients with a prolonged second stage of labor. *Obstetrics and Gynecology* 102(1):52-58.

Neuberg R (1995). *Obstetrics: a practical manual*. Oxford: Oxford University Press.

National Institute for Clinical Excellence (2001). *The use of electronic fetal monitoring*. London: NICE.

Nordstrom L, Achanna S, Naka K *et al* (2001). Fetal and maternal lactate increase during active second stage of labour. *British Journal of Obstetrics and Gynaecology* 108(3):263-8.

O'Connell M, Tetsis A, Lindow S (2001). The management of the second stage of labor. *International Journal of Gynecology and Obstetrics* 74(1):51-6.

Paine L, Tinker D (1992). The effect of maternal bearing-down efforts on arterial umbilical cord pH and length of the second stage of labor. *Journal of Nurse-Midwifery* 37(1):61-3.

Parahoo K (1997). *Nursing research: principles, process and issues*. Basingstoke: Palgrave Macmillan.

Parnell C, Langhoff-Ross J, Iversen R *et al* (1993). Pushing method in the expulsive phase of labor: a randomized trial. *Acta Obstetrica et Gynecologica Scandinavica* 72(1):31-5.

Perry L, Porter C (1979). Pushing technique and the duration of the second stage of labor. *The West Virginia Medical Journal* 75(2):32-4.

Petersen L, Besuner P (1997). Pushing techniques during labour: issues and controversies. *Journal of Obstetrics, Gynecologic and Neonatal Nursing* 26(6):719-26.

Polit D, Hungler B (1995). *Nursing research: principles and methods*. 5th ed. Philadelphia: JB Lippincott.

Roberts J (2002). The 'push' for evidence: management of the second stage. *Journal of Midwifery and Women's Health* 47(1):2-15.

Roberts JE, Goldstein SA, Gruener JS *et al* (1987). A descriptive analysis of involuntary bearing-down efforts during the expulsive phase of labour. *Journal of Obstetric, Gynecologic, and Neonatal Nursing* 16(1):48-55.

Roberts J, Woolley D (1996). A second look at the second stage of labour. *Journal of Obstetrics, Gynecologic, and Neonatal Nursing* 25(5):415-23.

Rooth G (1988). *Perinatal acid-base balance*. Sweden: Studentlitteratur.

Rossi M, Lindell S (1986). Maternal positions and pushing techniques in a nonprescriptive environment. *Journal of Obstetric, Gynecologic and Neonatal Nursing* 15(3):203-8.

Rushmer R (1947). Circulatory effects of three modifications of the Valsalva experiment. *American Heart Journal* 34(3):339-418.

Sampsel C, Hines S (1999). Spontaneous pushing during birth: relationship to perineal outcomes. *Journal of Nurse-Midwifery* 44(1):36-9.

Schultz K, Chalmers I, Hayes R *et al* (1995). Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 273(5):408-12.

Thomson A (1993). Pushing techniques in the second stage of labour. *Journal of Advanced Nursing* 18(2):171-7.

Westgate J, Garibaldi J, Greene K (1994). Umbilical cord blood analysis at delivery: a time for quality data. *British Journal of Obstetrics and Gynaecology* 101(12):1054-63.

Wood C, Ng KH, Hounslow D *et al* (1973). Time – an important variable in normal delivery. *Journal of Obstetrics and Gynaecology of the British Commonwealth* 80(4):295-300.

Woolley D, Roberts J (1995). Second stage pushing: a comparison of Valsalva-style with 'mini' pushing. *Journal of Perinatal Education* 4(4):37-43.

Yeates D, Roberts J (1984). A comparison of two bearing-down techniques during the second stage of labor. *Journal of Nurse-Midwifery* 29(1):3-11.

Bosomworth A, Bettany-Saltikov J. MIDIRS Midwifery Digest, vol 16, no 2, June 2006, pp 157-165.

Original article. © MIDIRS 2006.

See editor's note overleaf

Editor's comment

The above article is longer than the normal length we accept for regular original articles. The authors are to be congratulated as they have achieved no mean feat by reducing their original contribution to fit into the word limit. I think this piece of work is of interest on several levels. I decided to place this in the *Research and Education* section as I consider this aspect to be the one of priority; however, it is also about to be perhaps even higher on the clinical agenda as the *NICE Guidelines on normal labour* will be available for consultation in the forthcoming months and I hope this is an issue considered by these guidelines. With regard to research and education, I have two points to make; one is the requirement set by a University to insist on the inclusion of a set number of trials, in this case it was ten. While it is obviously reasonable to give guidance when undertaking comparative reviews, I would suggest that this does not need to be set at a defined number. My argument would be that it is the review itself which would determine the

number of studies included and that part of the education process is assisting the student to identify the inclusion and exclusion process. My second point is one for clinical practice. As identified by the authors, there would appear to be very little robust evidence for something that was, and still is in some areas, conventional or routine practice. In my view it is essential that we identify more often other areas where practice is so poorly underpinned by sound knowledge, and that midwives continue to become involved in challenging research approaches to midwifery care. We can then offer the midwives of the future a much better framework upon which to base their practice.

A few weeks after I was sent this article, I noted another study on this subject by Bloom *et al*. There is an abstract of this paper on page 234. Regretfully, I do not think it substantially increases our understanding but at least it shows an interest in this very neglected area, albeit by medical practitioners.